

Children's self-assessment of their school work in elementary school

Anastasia D. Elder*

Counseling and Educational Psychology, Mississippi State University, USA

(Received 19 June 2008; accepted 1 October 2008)

This study attempted to understand elementary students' self-assessment, that is, how they reflect on their own learning in school. Thirty-seven students (17 first graders and 20 fourth/fifth graders in a USA primary school) were interviewed about the kinds of standards against which they judge their school work and the sources of evidence they use in making self-assessments. Analysis of open-ended questions and Likert responses (1) highlighted that both younger and older students rely on evaluations provided by others, especially teachers, yet (2) pointed to developmental progression in which older students are increasing their use of self-standards and do consider multiple types of sources for their self-assessment.

Keywords: self-assessment; elementary school students; interviews; self-evaluation; academic work

Self-regulated learning is a fundamental goal of education (Paris and Paris 2001) especially in elementary grades when foundations for independent learning are established. Zimmerman's (1989) social cognitive model of self-regulated learning emphasises the importance of observing and evaluating one's own behaviour. Hence, a critical step for children to become self-regulated learners is to be engaged in appropriate self-assessment. Self-assessment refers to the process by which students come to gauge their level of performance and understanding. Self-assessment has been discussed in reference to students' metacognition (Andrade 1999; Paris and Paris 2001), self-efficacy beliefs (Chen 2006), and learning (van Kraayenoord and Paris 1997; Andrade 1999).

Related developmental research on children's self-perception of their competence has focused on young children's emotional reactions (Stipek, Recchia, and McClintic 1992), or have investigated students' appraisals of their ability or intelligence (e.g., Stipek and Gralinski 1996) or have linked general academic competence ratings to teacher and parent ratings (Herbert and Stipek 2005). Yet, there has been little research on children's self-assessment of their own learning and performance relative to specific academic tasks.

Towler and Broadfoot (1992) suggested that children's academic self-assessment should be a central goal for elementary school teachers who wish to foster motivated, responsible learners but there is not evidence about how or what children use in their judgements. Furthermore, other research emphasised that students' self-appraisals

*Email: aelder@colled.msstate.edu

are inaccurate and demonstrate their overconfidence (e.g., Dunning, Heath, and Suls 2004). Finding out more about the basis of their self-assessments – that is, the sources of information students use and criteria they rely on – would be helpful in understanding the basis for the inaccurate appraisals students have of their own learning. So, this study asked: from what sources do children gather information about their progress? What sorts of criteria or standards do students use in evaluating themselves? How does self-assessment change with age and schooling?

These questions were the focus of the present study which was designed to examine how children assess their progress in relation to school tasks. In particular, the focus was on aspects of the *sources* from which they create knowledge about themselves and the *criteria* against which they evaluate themselves. According to Stipek and MacIver (1989), young children relied on feedback from others and socially prescribed behaviour when they form judgements about their competence whereas older children used social comparisons and individual standards more often. Thus, it was hypothesised that young children will use the evaluations and tasks provided by others (teachers, parents, and peers) as sources of information and criteria of their own performance to a greater extent than older students who will use self-referenced processes (e.g., checking, proofreading) and standards (e.g., previous work, best effort) to gauge their accomplishments.

Method

Participants and context

Seventeen 6–7-year-old children (10 females and 7 males) and twenty 10–11-year-old children (10 females and 10 males) from three classrooms in a single public elementary school in the USA were interviewed. Participants represented a range of achievement levels, and all spoke English as their first language. None were classified for special educational services. All students in the three classrooms were asked to participate. Only those who assented and whose parents consented were eligible for participation.

The classrooms were recommended by the principal as having teachers who engaged students in reflective thought by using journals, portfolios, or other reflective tools. Hence, the students were in classrooms that practised some form of self-appraisal as part of their classroom activities. These activities were global in nature, that is, completing a sheet at the end of the week asking students to think about what made them feel proud, what they worked hardest on, what they liked most, what they liked least, and so on.

Interview protocol

Students were interviewed individually for approximately 30 minutes with four open-ended questions and 24 Likert-scale items (each with four response options) that asked children to discuss and rate their self-assessment in school. The semi-structured interview included questions about the *source* of evidence (e.g. people, activities) they use for evaluating their progress and the *criteria* they use in judging that their work is good. The open-ended interview items were inspired in part by others' work (van Kraayenoord and Paris 1997) and created by the author. The intent of the Likert items, in particular, was to represent the range of ways students may think about how they judge their work. Early pilot testing with third through

fifth grade students (not included as participants for this study) helped to refine the items.

Questions about sources

Two open-ended questions about sources for self-assessment began the interview: ‘Sometimes children know how well they are doing at school because of what other people do or say. What sorts of things do other people say about your schoolwork that helps you know how you are doing?’ and ‘Sometimes children know how well they are doing at school from the things they do in school. What sorts of activities in your class help tell you how you are doing?’ Students’ verbal responses were written as verbatim as possible by the investigator.

After generating responses to the open-ended questions, children verbally rated 12 items regarding the sources they use such as, ‘My teacher marks mistakes on my paper’, according to a 4-point scale on how frequently they used the source (4 = every day, 3 = every week, 2 = every month, 1 = hardly ever) and how helpful the source was for evaluating their learning and progress in school (4 = very helpful, 3 = somewhat helpful, 2 = not usually helpful, 1 = does not help at all). During this time in the interview when using the 4-point scale, a diagram was presented to students. This diagram pictured four ovals of decreasing size to help the children understand and use the stated scale accurately. All items were asked to the students verbally; students responded and notes (explanation, rating) were marked by the interviewer.

Questions about criteria

The open-ended questions about criteria for self-assessment included: ‘Imagine that you have just finished writing a story, but you have not shown it to anyone yet. You think it is really good and are proud of it. How would you know it is an excellent piece of work? How do you judge when your work is really good?’ Furthermore, students verbally responded to 12 Likert-scaled items regarding the criteria they use for evaluating themselves, such as, ‘I compare my work to what other students do’. This was accomplished in the same manner and with the same 4-point scale described above.

Coding open-ended responses

Coding categories for the open-ended questions were created for both the sources and criteria questions along two dimensions, person and task. That is, students’ responses were judged along dimensions of what people said (person) and along the work itself (task). The coding for the sources question was further subdivided into three levels – self (i.e. self-evaluation or self-directed projects), joint (evaluation with another or collaborative work), and other (others provide evaluation or tasks directed by others). See Table 1. In addition, all responses were coded, such that if a student named multiple ideas, each of these ideas was individually coded. Thus, a single student could be coded for one or more categories.

The coding of the criteria questions involved categorising responses according to person and task dimensions: naming self-standards (e.g., ‘I worked really hard’) versus others’ standards (e.g., ‘My teacher checks it’) for the person dimension or

meaningful text features (e.g., ‘My story makes sense, it is interesting’ or ‘I developed interesting characters’) versus using superficial text features (e.g., ‘Neat handwriting’) for the text dimension.

Analyses

Analyses of open-ended responses involved examination of proportions of students. In this study, the Fisher’s Exact test is applied when considering the significance of the association between developmental level (younger versus older) and whether the student mentioned a particular type of source or criterion (yes versus no). It is appropriate to use the Fisher’s Exact test when assessing the association between two nominal variables in a 2×2 contingency table (Siegel and Castellan 1988).

Results

Descriptive information of frequencies of coded responses for the open-ended responses is presented in Table 2 by each grade level. Grade-related changes in students’ self-assessment were found in the open-ended responses. Older students’ most frequent criteria for judging their work were self-standards. In fact, older

Table 1. Coding categories for sources of self-assessment.

Person	Self-evaluation: <i>‘I just know by looking over my work’</i>	Joint evaluation: <i>‘My dad helps me revise my reports’</i>	Others’ evaluation: <i>‘My teacher tells me I am making good progress’</i>
Task	Self-directed progress: <i>‘I am writing really well’</i>	Collaborative work: <i>‘Class discussions tell me what I need to improve on’</i>	Teacher-directed tasks: <i>‘Doing writer’s workshop’</i>

Table 2. Proportions of children who identify various sources and criteria by age.

	Younger	Older
Sources		
Person		
Self	.24	.10
Joint	.00	.10
Others	.82	.95
Task		
Self-directed	.06	.20
Collaborative work	.12	.25
Teacher directed	.41	.65
Criteria		
Person		
Self	.35	.75
Others	.41	.55
Task		
Meaningful	.18	.40
Superficial	.88	.35

students were more than twice as likely as younger students (75% to 35%) to name self-standards (e.g., ‘I read it over’ and ‘It contains things I like’) as criteria for self-assessment ($p = .017$ by Fisher’s Exact). Younger students (88%) most frequently named superficial text features (e.g., ‘I filled up the page’ and ‘I write my letters the right way’) as criteria for judging that their work was good. They were more than twice as likely as older students (35%) to name superficial text features ($p = .001$ by Fisher’s Exact). However, for both younger and older students, the most frequently named source of information for self-assessment was evaluation provided by others, 82% and 95%, respectively.

On average, older students offered a greater variety of sources ($p = .014$ by Fisher’s Exact) and criteria ($p = .038$ by Fisher’s Exact) for their self-assessments than younger students. One hundred percent of the older students were able to name more than one type of source whereas only 70% of younger students could. In naming criteria, 85% of older students (compared to 53% of younger students) provided responses in more than one type of category.

Children’s ratings of the helpfulness of various criteria in knowing how well they are doing are presented in Table 3. When entered in a single MANOVA, none of these showed grade-related or gender-based differences. However, the descriptive information suggested some interesting patterns. For older children, the highest rated criterion for telling them how they are doing was *trying hard* and *doing better than earlier work*. For the first graders, the highest rated criterion was *doing it without help*. Interestingly, for both younger and older students, the lowest rated criterion was *the task was easy to do*.

Discussion

This study provides preliminary information on the development of self-assessment and how the sources and criteria of children change over the course of their elementary years. Overall, findings point to a developmental progression in which (1) older students were more likely to name self-standards and meaningful text features as criteria for judging their work while younger students were more likely to refer to superficial text features, and (2) older students were eclectic in their use of sources

Table 3. Mean ratings of helpfulness of various self-assessment criteria by age group.

	Younger (n = 17)		Older (n = 20)	
	M	SD	M	SD
Try hard	3.44	0.86	3.73	0.47
Takes a long time	3.44	0.79	2.98	0.91
Is better than earlier work	3.35	0.72	3.70	0.55
Was difficult to do	3.00	1.05	2.83	0.80
Was easy to do	2.76	1.24	2.68	1.10
Did it without help	3.53	0.65	3.20	0.71
It is an important piece	3.50	0.83	3.08	1.13
Teacher writes comments	3.38	0.89	3.58	0.73
Parents talk to you about it	3.29	0.83	3.40	0.70
Your teacher gives your work a score	3.32	0.97	3.32	0.67
Other students comment on work	3.12	0.98	3.15	0.86
Is better than classmates	3.12	0.91	2.83	0.98

and criteria for judging themselves. Students of both age groups emphasised others (especially teachers and parents) in considering their sources and criteria for evaluating their work. So, older students appear not so much to replace others' evaluations for their own but are amassing a variety of ways to evaluate their work that includes varied sources. Such a development might point to the fact that external criteria have become internalised: students have taken feedback from others – teachers and parents – reviewed it and now refer to it as their own standards. Overall, these findings are consistent with other work that explores the developing self-awareness of children (e.g., Stipek and MacIver 1989).

This work points to the importance of understanding the sources and criteria students use. By understanding how students judge themselves, teachers gain insight into their metacognitive growth and, to some extent, students' motivations:

Students who are knowledgeable about their personal strengths and weaknesses and who understand the strategies that can enhance their performance may be better able to plan and monitor their work. They may also be able to make reasonable attributions for their own performance, in terms of their effort expended, task difficulty, and strategies used, so that they can retain a sense of self-efficacy and optimism when they encounter similar tasks in the future. (van Kraayenoord and Paris 1997, 534)

Furthermore, these findings point to ways that instruction and schooling has a role in students' self-understanding. These students were from classrooms that made reflection part of their schooling routine. Likely, a combination of instructional emphases on reflection and self-appraisal coupled with developmental capabilities served as impetus for changes noted. Other work has presented how instruction and school-wide approach can promote self-assessment (Towler and Broadfoot 1992).

Overall, research on children's self-assessment also has pragmatic implications for education. Current reforms which emphasise independent projects and portfolios for assessment require children to observe and evaluate their own learning and progress. Teachers are better able to help students when they know the standards they refer to or the kinds of information they seek and use while judging their progress on independent tasks. This study highlights the type of self-appraisal information that teachers can expect from their students. Teachers may be better able to help students when they know the sources and criteria students of various ages invoke when they gauge the quality of their work. These findings demonstrate that it is likely that young students just starting on their schooling have very authentic concerns for handwriting and their fine motor skills while students as young as fourth grade demonstrate ability to consider others' evaluations as well as turning to some internalised self-standards for evaluating their work. In doing so, they are moving toward more meaningful criteria for evaluating their own writing. Ultimately, by helping students reflect on themselves and their accomplishments, teachers can help students invest effort, interpret their progress, promote the internalisation of academic standards, and gain a sense of self-efficacy.

References

- Andrade, H.G. 1999. Student self-assessment: At the intersection of metacognition and authentic assessment. Paper presented at the annual meeting of the American Educational Research Association, April, in Montreal, Canada.

- Chen, P.P. 2006. Relationship between students' self-assessment of their capabilities and their teachers' judgments of students' capabilities in mathematics problem-solving. *Psychological Reports* 98: 765–78.
- Dunning, D., C. Heath, and J.M. Suls. 2004. Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest* 5, no. 3: 69–106.
- Herbert, J., and D. Stipek. 2005. The emergence of gender differences in children's perceptions of their academic competence. *Journal of Applied Developmental Psychology: An International Lifespan Journal* 26: 276–95.
- Paris, S.G., and A.H. Paris. 2001. Classroom applications of research on self-regulated learning. *Educational Psychologist* 36, no. 2: 89–101.
- Siegel, S., and N.J. Castellan. 1988. *Nonparametric statistics for the behavioral sciences*. 2nd ed. New York: McGraw-Hill.
- Stipek, D., and J.H. Gralinski. 1996. Children's beliefs about intelligence and school performance. *Journal of Educational Psychology* 88: 397–407.
- Stipek, D., and D. MacIver. 1989. Developmental change in children's assessment of intellectual competence. *Child Development* 60: 521–38.
- Stipek, D., S. Recchia, and S. McClintic. 1992. Self-evaluation in young children. *Monographs of the Society for Research in Child Development* 57, no. 1: 1–100.
- Towler, L., and P. Broadfoot. 1992. Self-assessment in the primary school. *Educational Review* 44: 137–52.
- Van Kraayenoord, C.E., and S.G. Paris. 1997. Australian students' self-appraisal of their work samples and academic progress. *Elementary School Journal* 97: 523–37.
- Zimmerman, B. 1989. A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology* 81: 329–39.

Copyright of Education 3-13 is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.